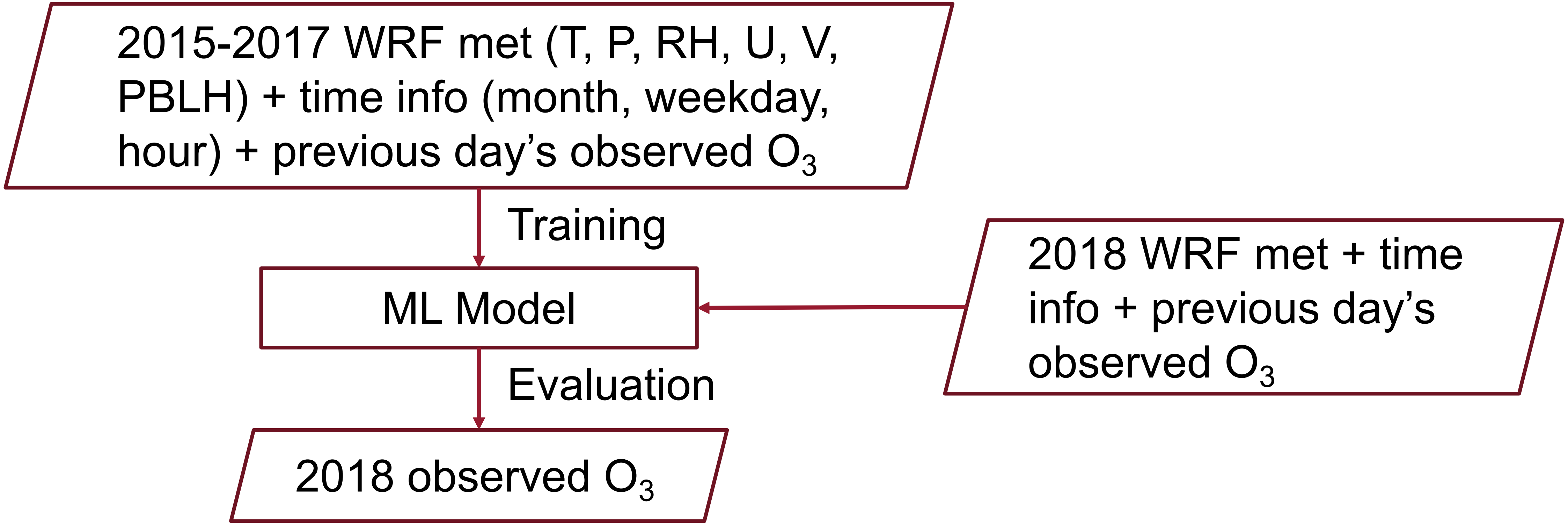
A Machine Learning Approach for Ozone Forecasting and its Application for Kennewick, WA

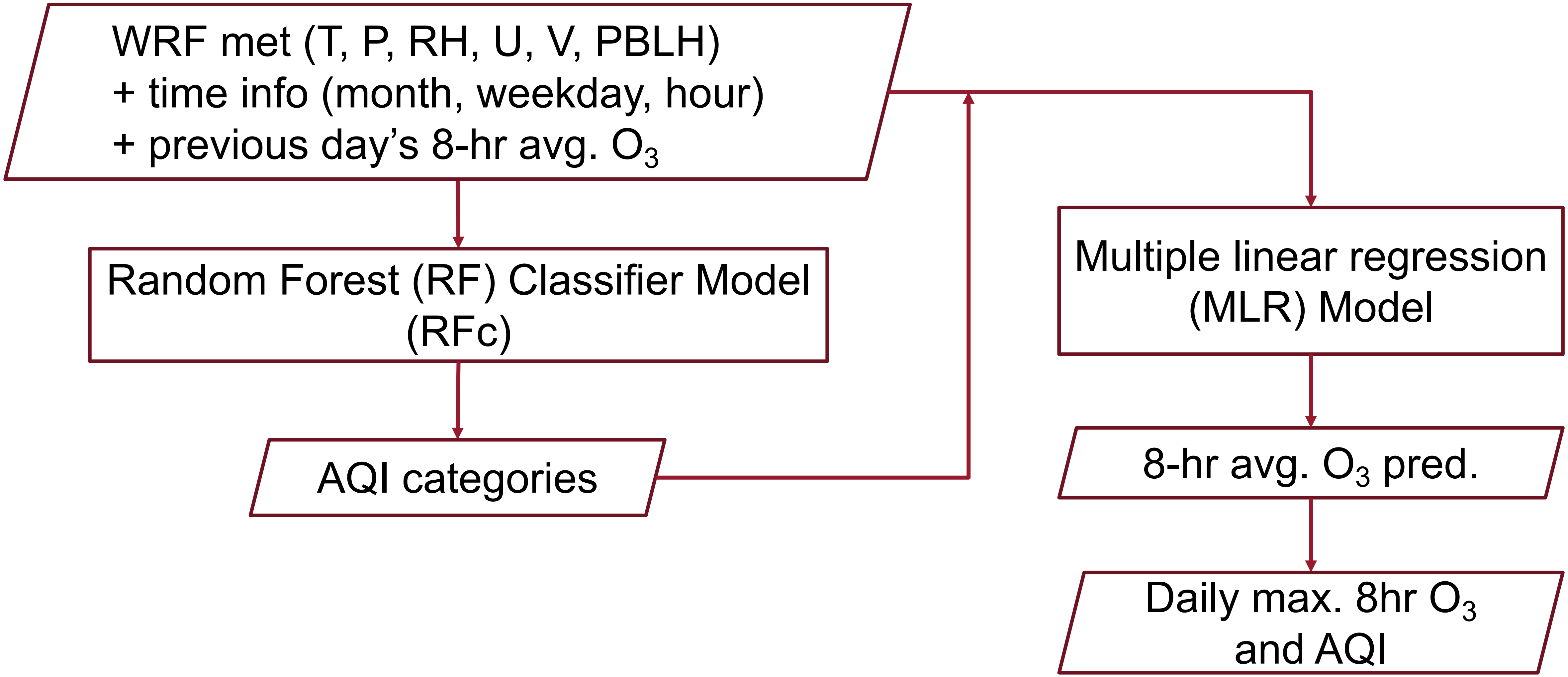
Kai Fan[1], Brian Lamb[1], Ranil Dhammapala[2], Ryan Lamastro[3], and Yunha Lee[1]

[1]Laboratory for Atmospheric Research, Civil and Environmental Engineering, Washington State University
[2]Washington State Department of Ecology
[3]State University of New York at New Paltz

Machine Learning (ML) Model Approach for the Kennewick Monitoring Site

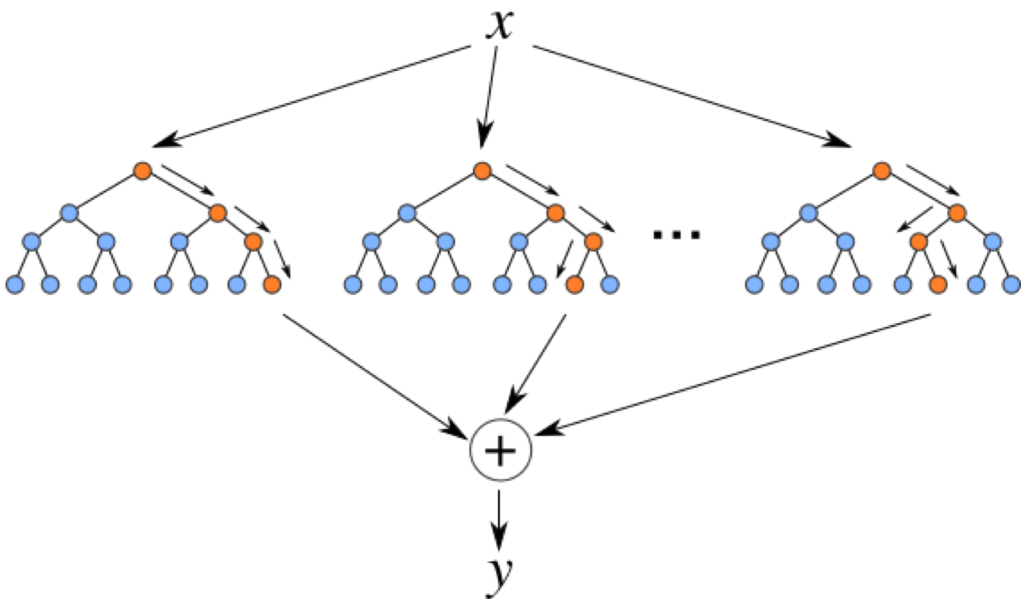


Machine Learning Model Framework 1: ML1
Combining Random Forest and Multiple Linear Regression methods



Random Forest (RF) classifier

- RF classifier is the consensus of many decision trees, which we use to predict the AQI categories.



* Image from https://blog.toadworld.com

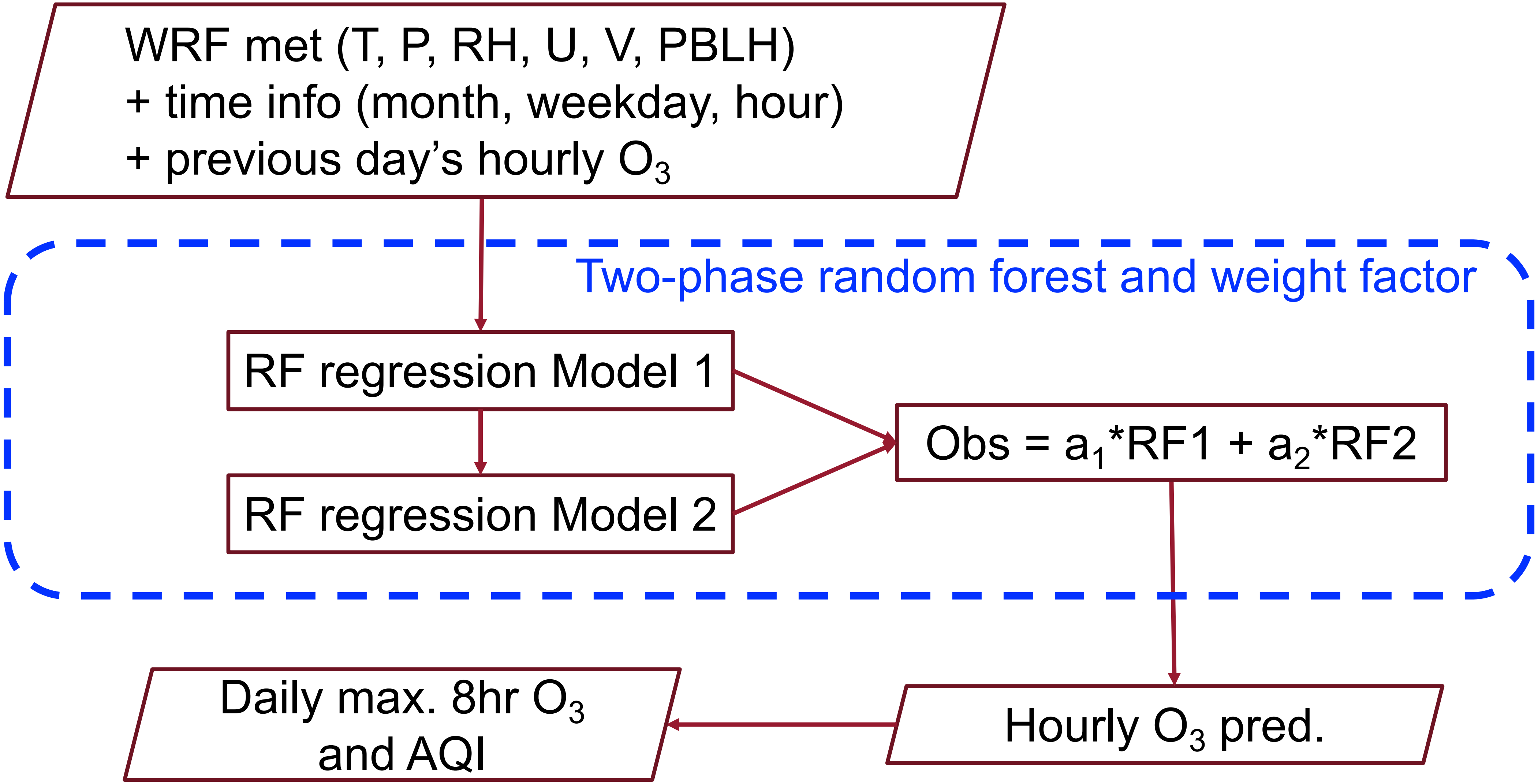
Multiple linear regression (MLR)

$$Y = a_0 + a_1X_1 + a_2X_2 + a_3X_3 + \ldots$$

- MLR approach is used to predict the 8-h average O_3, which shows good performance to predict high O_3 days.

Machine Learning Model Framework 2: ML2

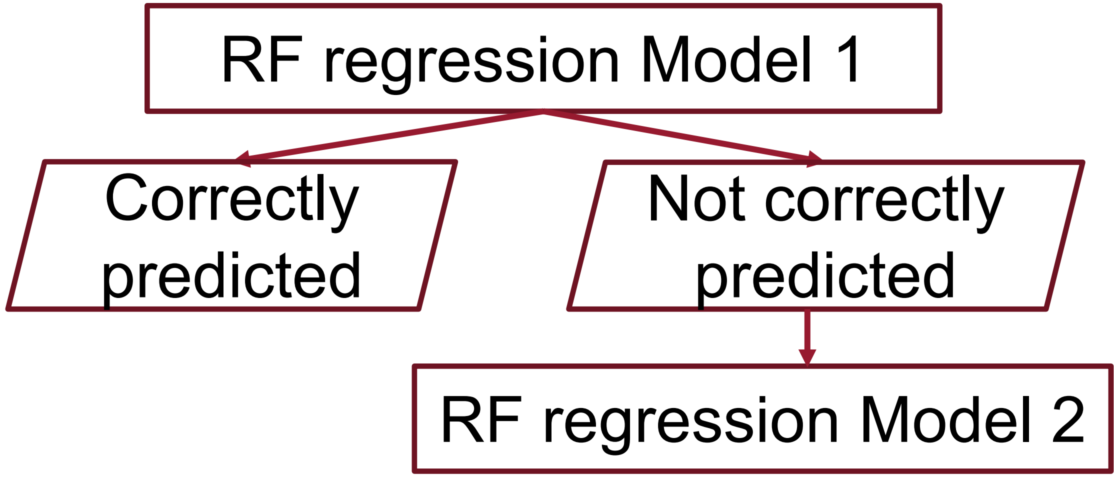
Two RF models weighted for optimal results



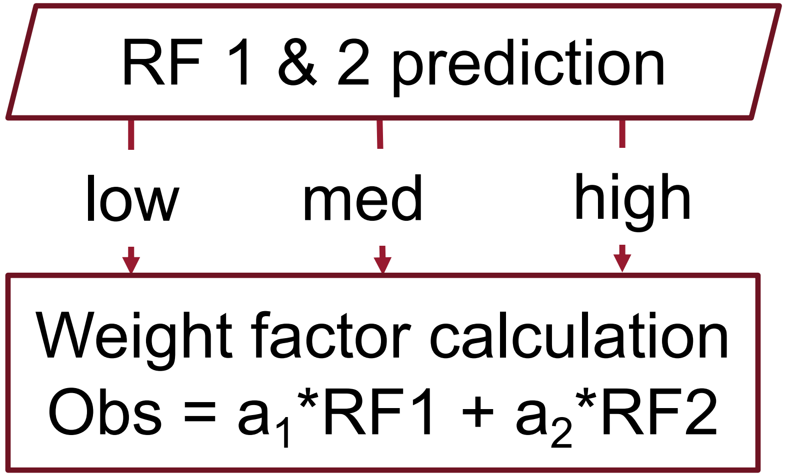
* Jiang, N., & Riley, M. L. (2015). Exploring the utility of the random forest method for forecasting ozone pollution in SYDNEY. *Journal of Environment Protection and Sustainable Development*, *1*(5), 245-254.

Two-phase random forest (RF)

- The first RF model can usually make right prediction for low O_3 events, and the second phase isolates the events incorrectly predicted to form a second training dataset.



- We separate the initial predicted mixing ratios to three categories and give three sets of weight to two phases. The weight of two models are based on a simple linear regression equation.



Forecast evaluation parameters

Heidke Skill Score (HSS)

- What is the accuracy of the forecast in predicting the correct category, relative to that of random chance?
- Range $-\infty$ to 1
- Perfect score = 1

The Hanssen-Kuiper Skill Score (KSS)

- How well does the model separate different categories?
- Range –1 to 1
- Perfect score = 1

Historical data summary

Year	Simulated days	# of days for each AQI			AQI > 2
		1	2	3	
2015	106	75	27	4	4%
2016	143	125	16	2	1%
2017	114	71	35	8	7%
2018	152	120	26	6	4%
Total	515	391	104	20	4%

More fires

ML1 Evaluation

		Observation											
		2015			2016			2017			2018		
		AQI 1	AQI 2	AQI 3	AQI 1	AQI 2	AQI 3	AQI 1	AQI 2	AQI 3	AQI 1	AQI 2	AQI 3
Model	AQI 1	57 (72)	4 (18)	0 (1)	91 (108)	3 (15)	0 (2)	50 (59)	6 (22)	0 (1)	93 (114)	0 (22)	0 (3)
	AQI 2	16 (2)	16 (6)	1 (1)	19 (2)	11 (1)	1 (0)	14 (1)	21 (9)	1 (6)	20 (0)	21 (2)	3 (2)
	AQI 3	1 (0)	5 (1)	2 (1)	0 (0)	2 (0)	1 (0)	1 (5)	5 (1)	6 (0)	1 (0)	3 (0)	2 (0)
HSS		0.45 (0.32)			0.42 (0.06)			0.53(0.24)			0.56(0.14)		
KSS		0.53 (0.25)			0.58 (0.04)			0.57(0.22)			0.72(0.09)		

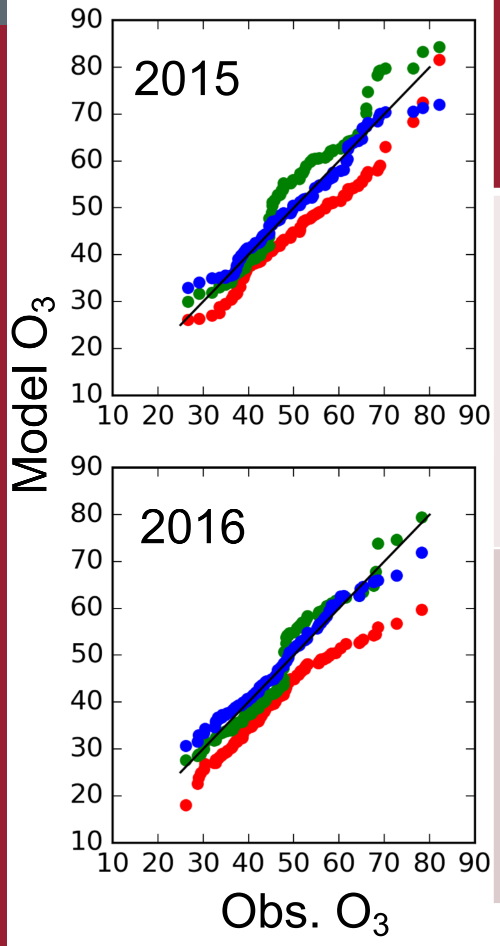
- Compared to AIRPACT, ML1 reduced underprediction but increased overprediction For high O_3 year 2017, ML1 performs better than AIRPACT.
- ML1 shows higher HSS and KSS than AIRPACT.

ML2 Evaluation

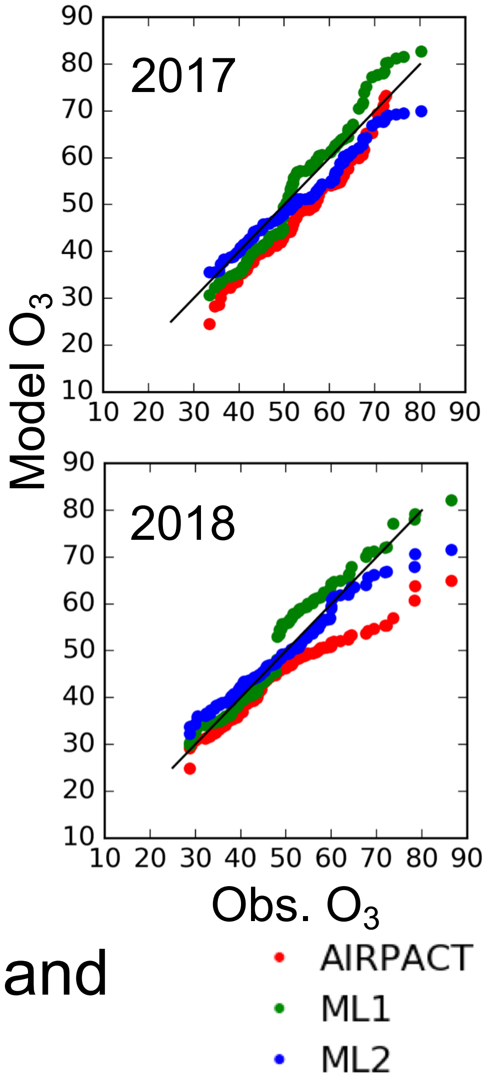
		Observation											
		2015			2016			2017			2018		
		AQI 1	AQI 2	AQI 3	AQI 1	AQI 2	AQI 3	AQI 1	AQI 2	AQI 3	AQI 1	AQI 2	AQI 3
Model	AQI 1	70 (73)	7 (18)	0 (1)	105 (112)	6 (16)	0 (2)	61 (58)	17 (23)	0 (1)	111 (116)	10 (23)	0 (3)
	AQI 2	5 (2)	17 (6)	3 (1)	8 (1)	11 (1)	1 (0)	3 (1)	16 (9)	7 (6)	5 (0)	14 (2)	5 (2)
	AQI 3	0 (0)	1 (1)	0 (1)	0 (0)	0 (0)	1 (0)	0 (5)	0 (1)	0 (0)	0 (0)	1 (0)	0 (0)
HSS		0.61 (0.32)			0.56 (0.07)			0.43(0.24)			0.54(0.14)		
KSS		0.59 (0.25)			0.59 (0.04)			0.39(0.22)			0.50(0.09)		

- ML2 shows higher HSS and KSS than AIRPACT. However, both AIRPACT and ML2 fail to predict the high ozone days in 2017.

ML1 vs. ML2

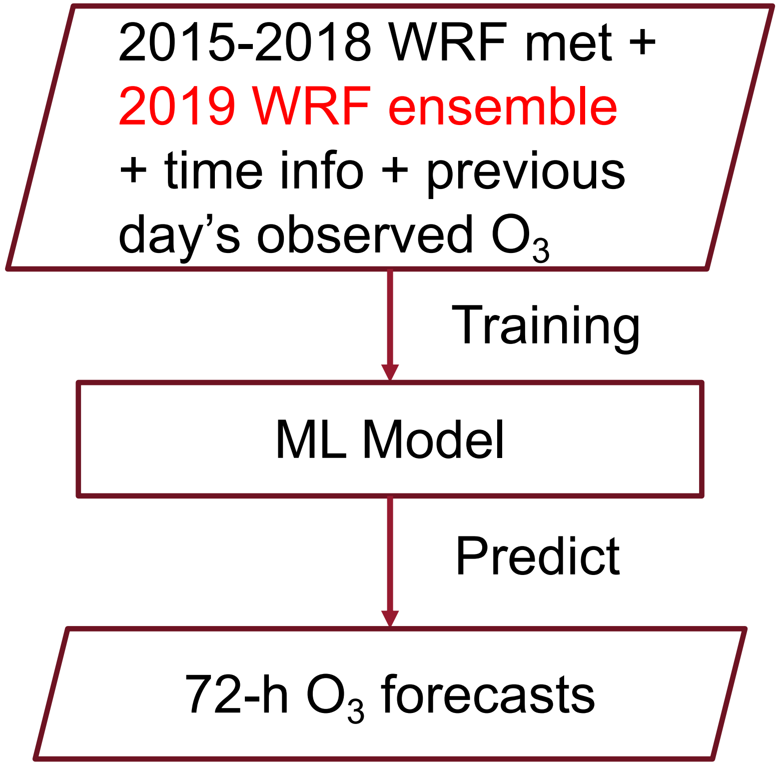


	Model	Year			
		2015	2016	2017	2018
HSS	ML1	0.45	0.42	0.53	0.56
	ML2	0.61	0.56	0.43	0.54
	AP5	0.32	0.07	0.24	0.14
KSS	ML1	0.53	0.58	0.57	0.72
	ML2	0.59	0.59	0.39	0.50
	AP5	0.25	0.04	0.22	0.09

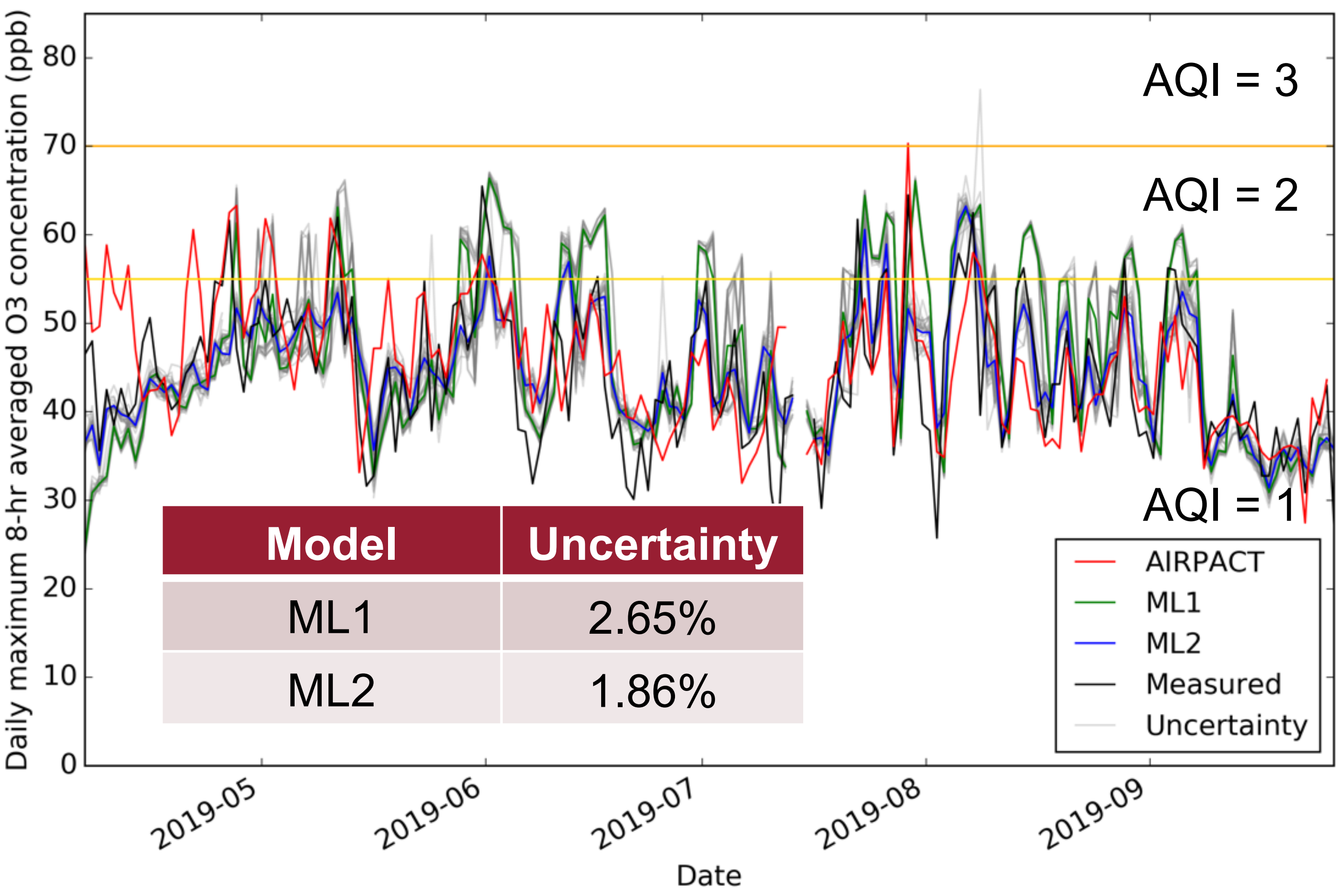


- ML1 and ML2 shows higher HSS and KSS than AIRPACT, and ML1 performs better in high ozone year (2017).
- Based on Q-Q plots, ML2 is usually closer to observations, especially below 70 ppb, and ML1 is better for high ozone events.

Tri-Cities Ozone "Ensemble" Forecast in 2019



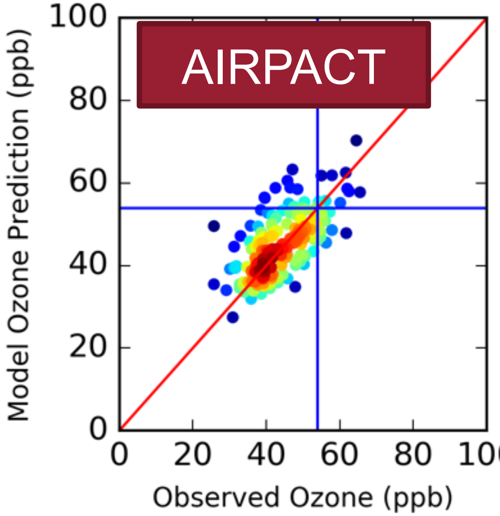
Time series of daily max. 8-h O_3

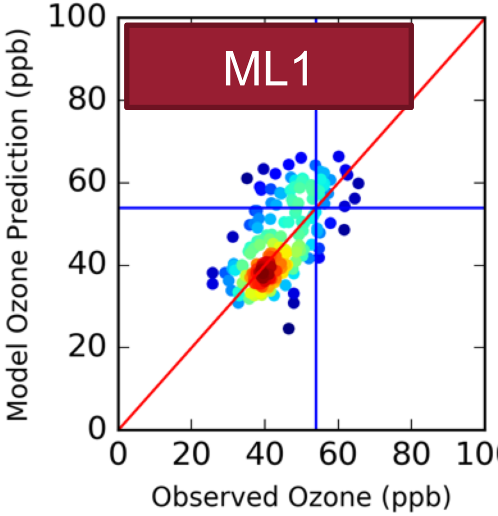


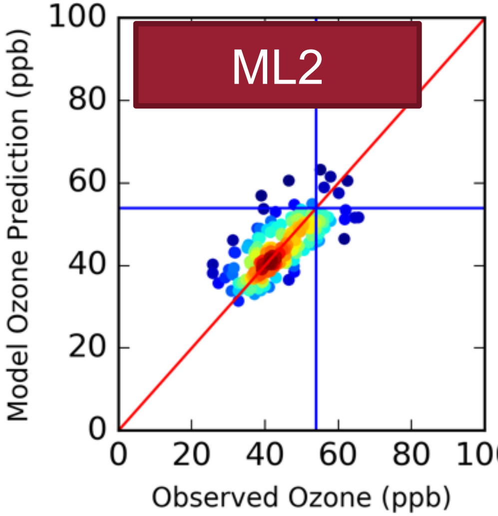
Model	Uncertainty
ML1	2.65%
ML2	1.86%

To get more data to train the model, we retrain our model everyday including previous day's measurements.

Tri-Cities Ozone (ensemble mean) Forecast in 2019







		Observation	
		AQI 1	AQI 2
AIRPACT	AQI 1	145	8
	AQI 2	10	9

		Observation	
		AQI 1	AQI 2
ML1	AQI 1	125	5
	AQI 2	30	12

		Observation	
		AQI 1	AQI 2
ML2	AQI 1	153	12
	AQI 2	2	5

ML2 performs the best to reduce false AQI2 days (in red cells).
Thus we chose ML2 to run our operational daily ozone forecasting for Kennewick.

Summary

- The ML1 model performed better in the high ozone year, but tended to overpredict ozone mix ratios.
- Both ML2 and AIRPACT missed some high ozone events. The Q-Q plot shows ML2 is closer to observations.
- Comparing HSS and KSS, both ML1 and ML2 performed better than AIRPACT. ML2 shows higher skilling scores except high ozone year.
- Our training dataset contains only a few high O_3 days, which makes it difficult to predict a high O_3 day using a ML approach. To overcome that issue, we updated the training dataset each day.
- We plan to apply our ML models to other cities that has a well-distributed AQI (Air Quality Index) values.

Thank you!